

## DATA TRAFFIC MANAGER AND METHOD THEREFOR

## BACKGROUND OF THE INVENTION

**[0001]** The present invention relates to a method and apparatus for data traffic management and, more particularly, to a time-spaced traffic management system for ATM data that uses per-class queues.

**[0002]** The growth and popularity of the Internet has led to tremendous amounts of data (voice, video, etc.) being transmitted from point to point and between servers. Multimedia and other high-bandwidth applications support Asynchronous Transfer Mode (ATM), which supports a wide variety of services and applications. An ATM network should be capable of providing differentiated Quality of Service (QoS) to these services and applications. Traffic Management protects the network and the end systems from congestion and promotes efficient use of network resources. Generic functions such as Connection Admission Control, Feedback Control, Usage Parameter Control, Cell Loss Priority Control, Traffic Shaping, Network Resource Management and Frame Discard are provided to facilitate efficient traffic management.

**[0003]** An ATM multiplexer connects different user devices (for example, DS1, DS3, T1/E1, Ethernet) then translates the protocols to ATM cell format and multiplexes the cells into a UNI interface (OC-3c or DS3 PLCP). The UNI interface may connect to the ATM switch. The combination of interfaces and services presents a difficult challenge to multiplexer implementation. The interface between the multiplexer and the network should support statistical multiplexing, traffic management, cell buffering, cell queuing, priority schemes, and fairness algorithms in accordance with data transmission standards. Traditional approaches require separate hardware and general purpose CPUs for each interface type and protocol,

which results in many individual designs and limited reuse of software.

**[0004]** Conventional shaping and scheduling algorithms used for ATM traffic management involve per-VC (Virtual Connection) queues. Such per-VC queues are processed with multi-level scheduling (two or more, for example, one for ATM classes and one for VCs) and shaping mechanisms. Typical shaping and scheduling algorithms include GCRA (Generic Cell Rate Algorithm), Priority Scheduling, WFQ, Scoreboard Scheduling, Fair Share Scheduling, Timestamp Scheduling, etc. However, per-VC queue processing requires a lot of computational resources and as such, is typically done in hardware.

**[0005]** Per-VC queuing and multi-level scheduling involves as many queues as the number of VCs and all of the queues must be scanned for each cell transmission. The larger the number of VCs, the larger the number of scans, making the processing complex and requiring an increasing amount of computing resources. Hence, some implementations compromise on the number of VCs that can be supported. Other implementations use two-dimensional timing chains that need frequent chain manipulations adding to computational overheads. Timestamp based scheduling mechanisms need a plurality of timers running at various time granularities making timer management complex. Some prioritized queue shaping mechanisms treat different classes of traffic alike. Some fair share schedulers perform round robin scheduling across VCs, however, do not distinguish between different classes of traffic.

**[0006]** One proposed method replaces per-VC shaping queues that must be processed during each cell-emission period with a sequence of shaping queues shared by a plurality of virtual circuits only one of which must be processed during each cell-emission period. However, queuing cells of different VCs/classes to the shaper queues and scheduling of the shaper queues is quite complex. A cell arriving late causes

designation of serving queues of all the VCs to be shifted, the impact of which is not clear. Also, for variable-bit-rate VCs, an additional overflow queue is required.

**[0007]** It would be desirable to provide a method and apparatus that performs efficient and cost-effective traffic management. It further would be desirable to have a traffic management process that does not cause defects.

#### BRIEF DESCRIPTION OF THE DRAWINGS

**[0008]** The following detailed description of preferred embodiments of the invention, will be better understood when read in conjunction with the appended drawings. For the purpose of illustrating the invention, there is shown in the drawings embodiments that are presently preferred. It should be understood, however, that the invention is not limited to the precise arrangement and instrumentalities shown. In the drawings:

**[0009]** FIG. 1 is block diagram of a data traffic manager in accordance with an embodiment of the present invention;

**[0010]** FIG. 2 is a block diagram of another data traffic manager in accordance with an embodiment of the present invention; and

**[0011]** FIG. 3 is a flow chart illustrating a modified generic cell rate algorithm of a data traffic manager in accordance with an embodiment of the present invention.

## DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

**[0012]** The detailed description set forth below in connection with the appended drawings is intended as a description of the presently preferred embodiment of the invention, and is not intended to represent the only form in which the present invention may be practiced. It is to be understood that the same or equivalent functions may be accomplished by different embodiments that are intended to be encompassed within the spirit and scope of the invention. As will be understood by those of skill in the art, the present invention can be applied to various packages and package types. In the drawings, like numerals are used to indicate like elements throughout.

**[0013]** The present invention is a data traffic manager that uses a time spaced round robin (TSRR) scheduler. The TSRR scheduler uses just per-class queues and avoids the use of per-VC queues and their associated complex processing. The TSRR algorithm is computationally efficient as it uses a modified GCRA and simple round robin mechanisms for processing the class queues. The TSRR scheduler significantly reduces the resources required, both processing and memory, thereby reducing the overall cost of ATM Traffic Management.

**[0014]** In one embodiment, the present invention is a data traffic manager having an enqueue engine connected to at least one port for receiving cells of data and a plurality of per-class shaper queues connected to the enqueue engine that receive respective cells of data from the enqueue engine. The shaper queues are spaced generally equally in time. The enqueue engine determines whether a cell of data is one of conformant and non-conformant. A conformant cell is queued to a currently served queue and a non-conformant cell is shaped based on a cell delay time. A time spaced round robin (TSRR) scheduler connected to the plurality of per-class queues

schedules cells of data from the shaper queues for transmission.

**[0015]** In another embodiment, the present invention is a data traffic manager that receives cell data of various class types from one or more data ports. The traffic manager includes a plurality of per-class queues connected to the one or more data ports, wherein at least one per-class queue is provided for each class of cell data and cell data from each class is queued in its respective queue. A first scheduler is connected to the plurality of per-class queues for scheduling the cell data and associated virtual connection (VC) traffic and Quality of Service (QoS) parameters. A shaper enqueue engine is connected to the first scheduler for managing the cells and their associated VC and QoS data. A plurality of per-class shaper queues are connected to the enqueue engine. The per-class shaper queues receive and queue respective cells and their associated data managed by the enqueue engine. At least one shaper queue is provided for each class. A second scheduler is connected to the plurality of per-class shaper queues for scheduling cells of data from the shaper queues for transmission. A timer is connected to the second scheduler and provides timing information thereto

**[0016]** In yet another embodiment, the present invention provides a method of managing ATM data traffic, comprising the steps of:

**[0017]** receiving ATM cell data from a data port;

**[0018]** analyzing the received data with an enqueue engine using a modified generic cell rate algorithm (GCRA);

**[0019]** placing the analyzed cell data into one of a plurality of per-class shaper queues; and

**[0020]** scheduling the cell data stored in the per-class shaper queues for transmission using a weighted/priority round robin algorithm.

**[0021]** Referring now to FIG. 1, a functional block diagram of a data traffic manager 10 in accordance with an embodiment of the present invention is shown. The data traffic manager 10 includes an enqueue engine 12, a plurality of per-class shaper queues 14, 16 and 18, a scheduler 20 and a timer 22. The enqueue engine 12 is connected to at least one port 24 for receiving cells of data. In the embodiment shown, the port comprises ATM Adaptation Layers (AAL) or other ATM ports. The enqueue engine 12 receives constant bit rate (CBR), real time-variable bit rate (rt-VBR) and non-real time variable bit rate (nrt-VBR) cells of data via the port 24. As understood by those of skill in the art, the difference between rt-VBR and nrt-VBR is the QoS parameters that they specify. rt-VBR specifies the same QoS parameters as CBR, *i.e.*, cell loss ratio, peak-to-peak cell delay variation and maximum cell transfer delay, while nrt-VBR specifies only CLR as a QoS parameter. The enqueue engine 12 may also receive unspecified bit rate (UBR) cells of data. The enqueue engine 12 determines whether a cell of data is either conformant or non-conformant using a GCRA or preferably, a modified GCRA, such as the modified GCRA shown in FIG. 3. A conformant data cell is queued to a currently served queue or the queue that is going to be served next if the current queue is full so that the conformant cell gets transmitted right away. For a non-conformant cell, the enqueue engine 12 determines whether the non-conformant cell has to be dropped or shaped based on how long the non-conformant cell has to be delayed to make it conformant and hence ready for transmission.

**[0022]** The plurality of per-class shaper queues 14, 16 and 18 are connected to the enqueue engine 12 and receive respective cells of data from the enqueue engine 12. The plurality of per-class shaper queues includes at least one CBR class queue 14, at least one rt-VBR class queue 16 and at least one non-real time variable bit rate nrt-VBR class

queue 18. Preferably, each class queue 14, 16 and 18 includes a plurality of such queues. For example, the CBR class queue 14 could comprise thirty-two (32) queues such that each CBR class queue can hold about thirty-two cells of data. The rt-VBR and nrt-VBR shaper queues 16 and 18 may include a similar number of similarly sized queues. The data traffic manager 10 preferably also includes a UBR class queue 26 for UBR data.

**[0023]** The scheduler 20 is connected to the plurality of per-class shaper queues 14, 16 and 18 and the UBR class queue 26 for scheduling cells of data from the shaper queues 14, 16, 18 and 26 for transmission. The scheduler 20 is preferably a time spaced round robin (TSRR) scheduler and in this embodiment, performs scheduling based on predetermined weight and/or priority values for each class, as well as timing information generated by the timer 22.

**[0024]** More particularly, for each class (CBR, rt-VBR and nrt-VBR), there are '**n**' shaper queues of size '**s**' (queue length - number of cells that a queue can hold), where **n** and **s** may be different for each class depending on granularity and a total number of queues available. The shaper queues 14, 16, 18 and 26 are spaced generally equally in time with '**Ts**' as a quantum of time where **Ts** = **s** \* **Tc** (cell time). The timer 22, which is connected to the scheduler 20, generates ticks for the cell time **Tc**. For example, for OC-3c, 1 cell time **Tc** is about 2.8  $\mu$ s. Put another way, each shaper queue within a class is serviced cyclically for **s** ticks in a cycle of **n\*s** ticks.

**[0025]** The shaper queues 14, 16, 18 and 26 are serviced cyclically for '**Ts**' duration and for each cell transmission only one queue from each class is scanned. For each tick **Tc** (cell time), W/P RR (Weighted/Priority Round-Robin) scheduler 20 scans currently served queues (one from each class) based on class weight/priority. In case of priority round robin, all the cells in the CBR queue 14 are scheduled before the

rt-VBR queue 16 and all cells in the rt-VBR queue 16 are scheduled before the nrt-VBR queue 18. Cells from the UBR class queue 26 are scheduled only when there are no cells to be scheduled from the other class queues 14, 16, and 18. In the case of weighted round robin, the CBR queue 14 is serviced  $W_{CBR}$  times, the rt-VBR queue 16 is serviced  $W_{rt-VBR}$  times, the nrt-VBR queue 18 is serviced  $W_{nrt-VBR}$  times and the UBR queue 26 is serviced  $W_u$  times, where  $W$  is a weight value.

**[0026]** Shaping is basically delaying the transmission of a cell by placing the cell in the appropriate shaper queue. While the scheduler 20 schedules cells from the shaper queues 14, 16, 18 and 26 for transmission, the enqueue engine 12 manages the shaper queues 14, 16 and 18 (but not the UBR queue 26). The time the cell has to be delayed is the difference between the current time and the time at which the cell should be transmitted. If this difference exceeds a predetermined value (referred to as  $maxCTD$  in FIG. 3), in one embodiment of the invention, the cell is immediately dropped. The shaper queue to be used for delaying the cell is determined based on the time difference and the queue that is currently being served. The cell is queued to the respective shaper queue based on queue thresholds. The cell is dropped or discarded if the shaper queue is above the non-conformant threshold of the class.

**[0027]** Referring now to FIG. 3, a flow chart of a modified GCRA used by the enqueue engine 12 is shown. As mentioned above, in this embodiment the enqueue engine 12 manages the shaper queues 14, 16 and 18 with a threshold tail drop mechanism and protects queue availability for conformant cells. Step 30 indicates the arrival of a cell  $k$  at a time  $t_a(k)$ . At step 32, the time at which the cell is to be sent or transmitted (Time\_to\_send) is determined as  $TAT - L$ , where  $TAT$  is a theoretical arrival time and  $L$  is a limit value.  $TAT$  can have an initial value of current time. For CBR,  $L$  is equal to

CDVT (Cell Delay Variation Tolerance) and for rt-VBR and nrt-VBR, L is equal to BT (burst tolerance) + CDVT, where, BT = (MBS - 1) \* (1/SCR - 1/PCR). MBS is maximum burst size; SCR is sustainable cell rate; and PCR is peak cell rate. At step 34, the calculated time to send is compared to the cell arrival time  $t_a(k)$ . If the time to send is not greater than the arrival time, then the cell is a conforming cell and the algorithm proceeds to step 36, where it is queued in the currently served queue for transmission. Step 38 follows step 36. At step 38, TAT is compared to  $t_a(k)$ . If TAT is less than  $t_a(k)$ , then at step 40, TAT is set to the value of  $t_a(k)$  and at step 42 TAT is incremented by I (the increment value I equals 1/PCR for CBR, and 1/SCR for rt-VBR and nrt-VBR). On the other hand, if TAT is not less than  $t_a(k)$ , then TAT is incremented by I at step 42. At step 34, if the time to send is greater than  $t_a(k)$ , then the algorithm proceeds to step 44 to determine the time that the cell has to be delayed. At step 44, the time to send less the arrival time  $t_a(k)$  is compared to a predetermined maximum value (maxCTD) and if the difference is greater than maxCTD, then the cell is dropped at step 48. If the difference is not greater than maxCTD, then the algorithm proceeds to step 46. At step 46, the shaper queue to which the cell is queued ( $Q_e$ ) is determined as  $Q_e = (Q_c + \text{Time\_to\_send} - t_a(k) + Q_c \text{time\_elapsed/s}) \bmod n$ . After the shaper queue  $Q_e$  is determined, TAT is incremented by I, at step 42 and then the next received data cell is processed.

**[0028]** Several implementations of the data traffic manager in accordance with the present invention are possible. However, the concepts and fundamental blocks like modified GCRA, enqueue engine, round-robin schedulers remain the same. Referring now to FIG. 2, another embodiment of a data manager 50 in accordance with the present invention is shown. The data manager 50 may be implemented using a RISC processor,

such as a Motorola C-Port Network processor. The data manager 50 has two levels of scheduling, one at ingress, before shaping, and another at egress, before transmission. At ingress, Weighted Round Robin (WRR) scheduling is used and at egress strict priority scheduling is used.

**[0029]** Like the data manager 10 in FIG. 1, the data manager 50 is connected to an AALs (ATM Adaptation Layers - upper layers of ATM) or other ATM ports to receive ATM cells. The received data cells are placed into respective class queues. Accordingly, the data manager includes a CBR class queue 52, an rt-VBR class queue 54, a nrt-VBR class queue 56 and an UBR class queue 58. A WRR scheduler 60 receives the data cells from the class queues 52, 54 and 56 based on weights assigned to each of the classes. In one embodiment, the CBR, rt-VBR and nrt-VBR class queues 52, 54 and 56 are assigned weights of 4, 3 and 1 respectively so that the CBR class queue 52 is serviced four out of eight times, the rt-VBR class queue 54 is serviced three out of eight times and the nrt-VBR class queue 56 is serviced one out of eight times.

**[0030]** The WRR scheduler 60 is connected to a shaper enqueue engine 62. The shaper enqueue engine 62 receives the cells dequeued by the WRR scheduler 60 and shapes them by placing the cells in various queues. It should be understood that the shaper enqueue engine 62 receives not just cells of data, but also their associated VC and QoS data. Like the enqueue engine 12, the enqueue engine 62 is connected to a plurality of CBR shaper queues 64, a plurality of rt-VBR shaper queues 66 and a plurality of nrt-VBR shaper queues 68. Also like the enqueue engine 12, the enqueue engine 62 maintains and manages the different sets of shaper queues for each class (except UBR class) with a simple tail drop mechanism and protects queue availability for conformant cells using thresholds. The enqueue engine 62 determines the shaper queue to be used for the cell with a modified GCRA (FIG. 3) as

described above. The shaper queues 64, 66 and 68 and the UBR class queue 58 are connected to a PRR scheduler 70 that schedules cells from the shaper queues 64, 66 and 68 and the UBR class queue 58 based on the class priority. Shaper queues within a class are serviced cyclically based on a timer tick  $T_c$  generated by the timer 22 and time spacing between queues  $(T_s = s * T_c)$  and across classes, a strict priority scheduling is used. That is, all cells in currently serviced CBR queue 64 are scheduled before rt-VBR queue 66 and all cells in currently served rt-VBR queue 66 are scheduled before nrt-VBR queue 68. Cells belonging to the UBR class and stored in the UBR class queue 58 are not shaped and are scheduled only when there are no cells in the other class queues.

**[0031]** In addition to the data traffic managers 10 and 50 discussed above, the present invention also provides a method of managing ATM cell data received at a data port. In the method, the received data is analyzed with an enqueue engine, like the enqueue engines 12 and 62 using a modified generic cell rate algorithm (GCRA) as shown in FIG. 3. The analyzed data is placed into one of a plurality of per-class shaper queues 14, 16 and 18 and the cell data stored in the per-class shaper queues is scheduled for transmission using a weighted/priority round robin algorithm, such as the one implemented by the scheduler 20. For an ATM traffic manager, the plurality of per-class shaper queues includes at least one CBR class queue, at least one rt-VBR class queue and at least one non-real time variable bit rate nrt-VBR class queue. Preferably however, for each class there are 'n' shaper queues of size 's', where n and s are different for each class depending on granularity and a total number of queues available. Each shaper queue 14, 16, 18 and 26 within a class is serviced cyclically for 's' ticks in a cycle of  $n * s$  ticks. **[0032]** A traffic manager in accordance with the present invention needs just one timer and scans only one queue per-

class during each cell time period, which makes the traffic manager computationally efficient, simple and scalable. As compared to prior art data traffic shapers, the present invention uses relatively simple to implement algorithms like the modified GCRA to determine the shaper queue for en-queuing cells and weighted or priority round robin scheduling for de-queuing cells from the shaper queues. Different sets of shaper queues for CBR, rt-VBR and nrt-VBR traffic classes are used and the scheduler dequeues cells from these sets of queues in weighted/prioritized fashion, which provides class isolation and prioritization. Further, Cell Delay Variation is controlled with 'n' shaper queues, 's' (queue length - number of cells that a queue can hold) and 'Ts' (time space between queues) for each class. It should also be noted that according to the present invention, the shaping decision is based on GCRA and cells arriving in burst are placed in appropriate shaping queues, as opposed to using an overflow queue.

**[0033]** The description of the preferred embodiments of the present invention have been presented for purposes of illustration and description, but are not intended to be exhaustive or to limit the invention to the forms disclosed. It will be appreciated by those skilled in the art that changes could be made to the embodiments described above without departing from the broad inventive concept thereof. For example, a data traffic manager can be used to manage other types of data besides ATM, such as variable size IP data. It is understood, therefore, that this invention is not limited to the particular embodiments disclosed, but covers modifications within the spirit and scope of the present invention as defined by the appended claims.